DOCUMENT RESUME

ED 332 502 FL 018 802

AUTHOR Williams, Jessica
TITLE Evaluating ITA Preparation Programs: Intensive versus Concurrent.
PUB DATE 90
NOTE 10p.; Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (24th, San Francisco, CA, March 6-10, 1990).
PUB TYPE Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Acculturation; Chinese; *Classroom Communication; Classroom Techniques; Communication Skills; English (Second Language); *Foreign Students; Higher Education; *Instructional Effectiveness; Language Tests; *Program Design; *Program Effectiveness; Program Evaluation; Summer Programs; *Teacher Education; *Teaching Assistants; Testing; Uncommonly Taught Languages
IDENTIFIERS *University of Illinois

ABSTRACT

A study evaluated the relative pedagogical and cost effectiveness of training programs for international teaching assistants (ITAs) that are conducted before and within the academic year. Both program types were established at the University of Illinois during the academic year 1989-90, with similar content, the same number of contact hours, the same instructor, and similar students in terms of first language background (Chinese) and range of English language proficiency. Twelve students participated in the 2-week summer course, and 10 in the 10-week concurrent course. Data were gathered on students' pre- and post-test scores on a language test and on an instrument designed specifically to evaluate the teaching and language skills of ITAs, and departmental evaluations were obtained. Follow-up observations of course participants were also made. Based on these results and on impressionistic information, it was concluded that given limited resources, the best solution is to provide a concurrent course focusing on pedagogical effectiveness, including language and interaction skills and classroom culture and management. A less intensive summer course might be added, more geared toward social and cultural orientation, to feed into the concurrent course, which could concentrate on classroom communication. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Williams
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED332502

Evaluating ITA preparation programs: intensive versus concurrent

Recent surveys of ITA preparation programs around the country indicate that there is a trend away from concurrent, semester or quarter-long programs and toward intensive programs which take place prior to the academic year. While there are a number of arguments to explain this trend, there seems to be little empirical basis for establishing one or the other program type. This study was an attempt to make a preliminary determination of the relative pedagogical and cost effectiveness of the two program types.

Let me briefly review some of the arguments behind the establishment of both program models. First, the pros and cons of the intensive pre-academic year model. It has been argued that this is the time when intervention is the most crucial. ITAs need orientation to the university, the United States, the educational system, their new duties, as well as language and pedagogical training. This is clearly the time during which they are most at sea. It is also probably the only time during which ITA programmers can have the ITAs' undivided attention. Later in the term, their focus is on their content studies, their research, and one hopes, on their teaching. Prior to the term, the main competition is probably housing and other setttling in concerns. Cost effectiveness is somewhat more difficult to estimate because these intensive summer programs run anywhere from a one-week orientation to a two-month preparation course. In either case, however, room, board and stipends are a major concern. Very few of these students have the resources to support themselves beyond the academic year, especially when they first arrive. That means somebody else has to do it, usually either the university or the individual departments. At those universities where a concurrent program is already in place, it may be difficult to convince these centers of responsibility that they need to fork over more money when the problem is already being "taken care of." Very good documentation is therefore would be needed to persuade them that the intensive format is the more effective. Another drawback of this model is that the participants in such a program will have relatively little experience to draw on. They probably know very few people; their only major source of input will be their instructor and perhaps a few department members who

FL018802

BEST COPY AVAILABLE

are not on vacation. Their primary opportunity for interaction will be in the ITA classroom. Finally, from an administrative point of view, it may be more difficult to control for the proficiency level of students who shows up to participate in a pre-term intensive course. Of course, most ITAs will have a TOEFL score, but very few will have taken the TSE, and even for those who do, this information may not always be as helpful as one might wish. Telephone interviews are sometimes used for initial screening, but these are not without problems either (Fox, Berns and Sudano 1989).

Now I will turn to the advantages and disadvantages of the second type, the concurrent program. For better or for worse, within this type of format, some of the "orientation" may have already been accomplished, albeit of the sink or swim variety. As a result, the curriculum can be planned somewhat differently, with perhaps more emphasis on the pedagogical issues and language training, and less on survival. However, it is likely that these ITAs are taking the preparation course on top of an already heavy schedule and it is also likely that fulfilling the requirements of the class are not their first priority. It is therefore impossible to make the same sorts of demands, in terms of workload and commitment, as in a summer intensive course. Many ITAs and their advisors might prefer to have the training "over and done with" before they have to launch into their new responsibilities. On the other hand, the students may be able to bring more insights from their own to the classrooms, having experienced univeristy life first hand. The instructor may be able to tailor the course to the needs and concerns of the ITAs, whereas in the intensive course, we can only predict problem areas on the basis of past experience and some judicious guesswork. There is perhaps less "telling" and more interaction in the concurrent course. Indeed, one of the advantages of the a this type of course is the degree of input and interaction opportunities which are available to the ITAs, in comparison to the intensive course. Whether or not they take advantage of these opportunities is another question altogether, of course. Encouragement or enforcement can be built in as part of the concurrent course curriculum. Finally, there are administrative and financial advantages to the concurrent course. The bottom line is that the ITAs are here already. No additional support needs to be provided outside of instruction. At most, the generous department has to relieve the ITA of one course, so that he or she can devote more time to the ITA program. If a department has seen a pattern of good

results, they are usually more than willing to do this, since it avoids headaches later on. For the ITA program director, this format can be easier since the ITA is on the spot, not some disembodied voice or a score on a piece of paper. It is far easier to make judgments about what kind of training is needed, if any.

All of these arguments, however, are either based on intuitions or on non-pedagogical concerns. The decision should be founded on one question: which format works better. Of course, this question is not as simple as it sounds. First, there is the problems of apples an oranges. How intensive is the intensive program--one week or eight? Is the concurrent program one hour a week for a year or eight hours a week for a quarter? What are the variables to be investigated and equally important, how can we measure relative effectiveness?

In an effort to determine which of these program types is the more effective, both types of programs were set up at the University of Illinois-Chicago during the academic year 1989-90. The concurrent program had already been established in compliance with a 1986 state law. The intensive program was supported by a grant from the Fund for the Improvement of Post-secondary Education, expressly to determine the answers to some of these questions. The two programs had similar content, the same number of contact hours, the same instructor, and similar students, in terms of first language background (all were Chinese) and English proficiency. I should qualify this latter statement by saying that not all of the students showed similar language proficiency but a similar range of ability was found in the two groups. All were recent arrivals in the United States. Some had had teaching experience in China, others had not. Twelve ITAs participated in the summer course, which lasted two weeks, five hours a day, four days a week. Ten ITAs participated in the concurrent course which ran for ten weeks, four hours per week. The original goal of the study was to evaluate the effectiveness of the two programs by using: (1) pre- and post-test scores on the SPEAK test, (2) pre-and post-test scores on an instrument designed specifically for evaluating the teaching and language skills of ITAs, (3) undergraduate student evaluations (4) departmental evaluations, (5) the grades of the ITAs' undergraduate students, controlling for ACT scores. As it turned out, it was only possible to gather quantitave data on (1) and (2), and

qualitative data on (4). Additional information was collected in the form of follow-up onservations of course participants.

I will first outline the quantified results we did get in comparing the two types of programs. Second, I will explain some of the problems we ran into in doing this kind of evaluation study and why we were unable to collect all of the data we had planned to. Third, I will discuss some of the reactions of the ITAs who participated in the two courses, the departments and well as some of my own, as I conducted follow-up observations. Finally, I will make some suggestions as to where I think we should go with these results regarding choices among program types.

At the beginning and at the end of each of the two courses, students took the SPEAK test and made a brief, taped oral presentation, similar to what they would have to do in an introductory recitation or lab session. In an effort to avoid the worst of practice effect, two different versions of the SPEAK were used and the students were required to "teach" different topics on the two occasions. Admittedly, when they first arrive, the ITAs are bewildered and often bumbe their way through both tasks in this pre-test. However, I am assuming that participants in the two groups were equally befuddled and I only report on the comparison of gains. One could certainly make the argument that the demonstration of gains in the first place has more to do with a growing familiarity with the tasks than with actual learning. I would argue that this too, is progress and while we may not be measuring what we think we are, it is some sort of improvement. And again, I stress than this is not a pre-test/post-test measure; it is a comparison of net gains across program types.

Before I report the results, I want to discuss briefly the instrument we used to measure the teaching portion of the evaluation. We have been working on this for a year or so, piloting it in various forms and with various judges. Basically, it includes an evaluation of language, interactional and presentation skills. Judges are asked to evaluate the frequency of various behaviors. This is in contrast to many other instruments which ask judges to rank the quality or sufficiency of these behaviors. It was felt that the frequency assessment was a more objective one. After a number of versions, we found this one to be the most consistent. Judges evaluate these behaviors on a one to three scale, ranging from "often" to "rarely" on such questions as "Does the speaker make grammatical errors in simple

sentences?" and "Does the speaker teach with his/her back to the class?" etc. An initial questionnaire with a five-point scale containing the extremes "always" and "never" was rejected after we discovered that these alternatives were rarely chosen and hence did not discriminate well among subjects. The instrument contains 35 questions, each with a possible score of 3. The total scores were divided by 35 to come up with an overall score of between 0 and 3. It was piloted with native speaker and non-native speaker subjects; all of the native speakers received a score of 2.87 or above. We tested the instrument with both undergraduate students and ESL instructors and found that the scoring of the two groups was relatively consistent, with the ESL teachers giving slightly lower scores on the language portion of the assessement. In general, however, this pilot phase made us confident that our own judgments, as TESL professionals, were not far from those given by undergraduates who, afterall, are the ultimate beneficiaries of these services.

Before trying to determine which of the two program types resulted in greater gains, I first wanted to be sure that the gains in fact were significant. The Wilcoxon signed ranks test indicated that the gains were in fact real for both program types and for both the SPEAK and the video presentation. The gains in the summer were shown to be significant at $p<.005$ (N-8) and in the concurrent course at $p<.025$ (N-6). These were, of course, encouraging findings, but they say nothing about the comparative effectiveness of the two programs.

Net gains on two assessement measures

	Intensive	Concurrent	
Lesson Presentation			
mean	.14	.33	
s.d.	.06	.08	U=0 $p<.005$
SPEAK			
mean	27.50	23.33	
s.d.	15.81	12.11	U=19 n.s.

Your handout shows the figures for both tasks for the two groups of ITAs. The results of the Mann-Whitney U-test indicate that the gains in the

SPEAK test were not significantly different across program type. On the teaching task, however, the ITAs who participated in the concurrent course posted significantly greater gains than those in the intensive course. These results should be interpreted with caution, however. Several of the ITAs in the concurrent course had already been assigned teaching duties during the time they were taking the course. These also tended to be the subjects who improved the most on this task. It may therefore have little to do with the course itself and simply prove the old adage that one learns by doing. One might also ask whether the short time span of the intensive course had an important effect on the score. Would we have found the same result with a course which had lasted four weeks instead of two, but with the same number of hours? I mention all of these possiblilities to stress that these results are very preliminary and much more would need to be done in order to dctemine the best format for ITA preparation courses.

These are not the only problems and questions which we encountered in conducting this research. As I mentioned earlier, we had originally hoped to collect information on the success of these ITAs in the classroom after they had completed the preparation courses. We were even hoping to gather data on the success of their undergraduate students. This was an area of particular interest to the granting agency. Between Tiananmen square and our own enduring naiveté, we were unsuccessful in these areas. First of all, we had far fewer subjects than we wished and as a result were able to be far less selective in assigning ITAs to courses. It was essentially first come-first serve and the ITAs had a much greater range of English proficiency than we had planned. In addition, as luck would have it, we got the bulk of one department's ITAs in one course and the bulk of another's in the second course. We had hoped to assess ITA effectiveness across lecture sections in individual courses. As it turned out, however, there was not a single matched pair of say, one ITA from the intensive course and one from the concurrent course assigned to TA under the same lecturer in the same course. As a result, data on student evaluation and success were not gathered. In many cases, the ITA "graduates" were not assigned to teaching posts at all, and instead were given grading duties or became RAs at the last minute. What with the usual attrition for other reasons, the bottom pretty much fell out of our plans for long-term evaluation.

In an effort to flesh out the assessment, I have conducted observations of all of the ITAs in the two courses who had been given teaching assignments and interviewed the graduate chairs of the departments with the largest number of participants, namely, Physics, Chemistry and Mathematics. Clearly, these are both somewhat impressionistic and I must admit that I could not detect any obvious difference between the participants of the two courses during my observations. I did try, however, to follow-up on the same categories which we had selected for testing as part of the assessment instrument. Again, anecdotally, the biggest changes which I detected were in classroom management, presentation and interaction skills. No matter how many times we had gone over the pronunciation of individual key words used in their field, "trigonometric" and "parabola" remained garbled. On the other hand, they made conscious effort to frame and segment their presentations, to check periodically for comprehension and in general seemed more at ease in front to their classes than the many ITAs I have observed over the years who have had no explicit training.

In talking to their advisors regarding the summer intensive course, I discovered that again, improvement in language skills was not what they had noticed. Often, they said that they could not quite put their finger on it, but that these entering ITAs seemed somehow "ready to go" from the first day. They were not bewildered and stymied by the many administrative details with which all new TAs have to cope. They exuded more confidence are seemed more willing to interact and take risks. I realize that these reactions, both mine and other faculty's, is subjective at best. However, I am using them to go out on a limb and make a recommendation for choosing among ITA program types.

Before I do that, however, I should say a few words about the relative cost of the two programs. This is somewhat difficult to assess in a way which would be useful to other universities, since there seems to be wide variation in where ITA programs are housed and who pays for them. I would guess though, that for most universities, the summer program is going to be more expensive, mostly because someone, either the individual departments or the school or university, is going to have to foot the bill for their upkeep. At some universities, including mine, full services are not available during the latter part of the summer, so it may be necessary to

make special arrangements to open buildings and to provide food and accommodation. I ended up having a restaurant in Chinatown cater the meals, since there was no food service during this period. In short, there may be a lot of financial and administrative hassles associated with a summer program. The cost of hiring an instructor, of materials and other indidentals remains relatively constant. Thus, from the cost effectiveness angle as well, it appears that the concurrent program is the better bet.

Based on all of these bits and pieces of information then, some quasi-empirical, some unabashedly impressionistic, these are the conclusions that I have come to. Given limited resources, the best solution is to provide a concurrent course, focussing on pedagogical effectiveness, understood in the broadest sense to include language and interactions skills, classroom culture and management etc. My next addition would be a summer course, but one which would be less intensive than the one which we ran and would focus on a general orientation to the country, to the university, to American education and classroom culture. This would then feed into the concurrent course, which would then be able to focus less on socio-cultural orientation and more on the immediate issue of classroom communciation. Finally, and this has nothing to do with the project I am reporting on today, I think it is crucial to include a follow-up component. I know that many larger universities already do this, but it is something of a luxury. We have found that we have been able to increase our effectiveness by including content faculty and NS TAs. We have just added a course which has an observation-followed-by-feedback format. An ESL ITA trainer and a NS content TA jointly observe the new ITA in his or her own class. We then gather those from the same department with similar problems in small groups--generally two or threes--to offer suggestions and discuss problems they have had. I don't have any formal evaluation data on this component since, as I say, we have just started it. However, since I am the ESL person, I can say that having the content person along has been a tremendous help and extremely efficient. There are many things which simply fly right by me and and picked up immediately by the content TA. I believe that if, in the best of all possible worlds, we could offer all three of these components, we would bhave an extremely effective program, at least for those ITAs who enter with basic language skills (which, of course, they do not always do).

Clearly, this is just a pilot study, but as far as I know there is little empirical evidence of this kind. I hope that it will be helpful to those trying to decide which type of program to establish.

References

Fox, W., M. Berns and G. Sudano. (1989). *Assessing the communicative competence of ITAs: the telephone interview.* Paper presented at the 2nd National Conference on the Training and Employment of Teaching Assistants, Seattle, WA.